Virtual Shadows in Mixed Reality Environment Using Flashlight-like Devices

Takeshi Naemura,[*1]

Takuya Nitta[*2], Atsushi Mimura[*2], and Hiroshi Harashima[*2]

[*1]   [*2]   [*2]   [*2]

Abstract – We propose the concepts of Virtual Light and Virtual Shadow with the aim of achieving a Mixed Reality Environment focused on shadows. In this proposal, we divide the concept of Virtual Shadow into four categories, and among them, implement four types of interactive applications: (a) real to virtual shadow for rigid objects, (b) real to virtual shadow for non-rigid objects, (c) image-based virtual to virtual shadow, and (d) virtual to real shadow. In these applications, we can see the shadow of a real object projected onto the virtual world and vice versa. These proposed concepts and applications should contribute to the realization of a Mixed Reality Environment that provides a novel sense of visual interaction.

Keywords : virtual light, virtual shadow, shading and shadowing, flashlight interaction, seamless marging of the real and virtual worlds

1. Introduction

The seamless merging of the real world that we live in and the virtual world constructed within a computer is a challenging topic in current Virtual Reality research. Technology that superimposes and displays computer generated images in the real world is called "Augmented Reality," and conversely, technology that enhances the virtual world using real world data is called "Augmented Virtuality." In either case, the objective is to merge the real world and virtual world. From this point of view, the concept of "Mixed Reality" that includes both Augmented Reality and Augmented Virtuality has been proposed [8].

Three key issues in achieving Mixed Reality are geometry, illumination, and time [9]. Needless to say, accurate alignment of spatial relationships in the real world and virtual world is extremely important (consistency of geometry). Likewise, shade and shadow in both worlds must match to achieve a natural merge (consistency of illumination). Finally, movement in the two worlds must be synchronized to facilitate smooth interaction (consistency of time).

Here, we focus on the second issue, consistency of illumination, and propose several techniques for interactively projecting shadows of real/virtual objects onto the real/virtual worlds in a Mixed Reality Environment. These proposals are implemented so that we can move and rotate both the object and a light source to see the deformation of shadows interactively. When the object is a real one, we can handle it by our one hand and project its shadow onto our surroundings from a flashlight-like device (called "Virtual Light") in our other hand as desired. When the object is a virtual one, we may need a kind of interface or equipment to handle it, but we can still project its shadow interactively by the flashlight-like device. In this way, the user experiences a Mixed Reality World that seamlessly merges the real and the virtual.

2. Related Works

As an example of interaction using a flashlight-like device (Virtual Light), Inoue [2] has proposed an application that makes a virtual object transparent and its internal structure visible. Lutz et al. [4] have applied a tracked flashlight to control a virtual light source in the CAVE environment. We can use and
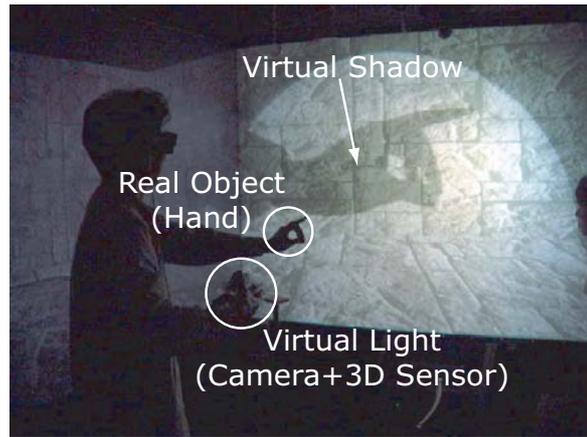
*1 Graphics Lab., Stanford University, USA
*2 Dept. Inform. & Commun. Eng., The Univ. of Tokyo
*1
*2

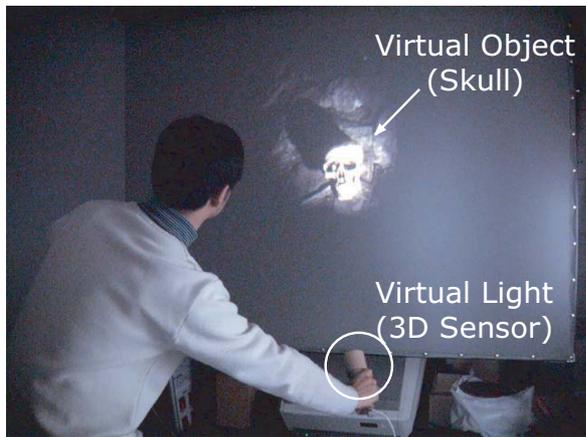(a) RV Shadow for rigid objects (Sect.5.1): The shape of a real object (a toy cat) is measured in advance and utilized for displaying its shadow onto a virtual wall.

(b) RV Shadow for non-rigid objects (Sect.5.2): The shape of a real object (a hand) is measured in real time, and utilized for displaying its shadow onto a virtual wall.

(c) Image-Based VV Shadow (Sect.5.3): While the shape of a virtual object (a skull) is unknown, the image-based method makes it possible to display its shadow onto a virtual wall.

(d) VR Shadow (Sect.5.4): The shadow of a virtual object (a teapot) is projected onto the real world. We can see just the shadow, and we need an additional display (See-through HMD) to see the teapot itself.

1

Fig. 1   Four examples of Virtual Shadow Interaction.

expand upon the concept of Virtual Light as a means of merging the real and virtual worlds (see Fig. 1).

Even if the shape of a real object is unknown, there has also been research on treating real objects as virtual objects by preparing a large number of images. Katayama et al. [3], for example, have proposed a technique for selecting appropriate shading according to lighting conditions in the virtual world from a large number of images of a real object previously captured under various lighting conditions, and a technique for synthesizing a shadow from the silhouette of the object. Wong et al. [11], [12], mean-

while, have applied the idea of a bi-directional reflectance distribution function (BRDF) to the large number of images and have proposed a technique for synthesizing an appropriate image according to lighting conditions in the virtual world. In addition, Nishino et al. [6] have used a coarse-shaped model to describe shading for the surface of an object. We can introduce the flashlight-like device to the image-based methods like the ones above for the purpose of Virtual Shadow Interaction (see Fig. 1(c) and Sect. 5. 3) [5].

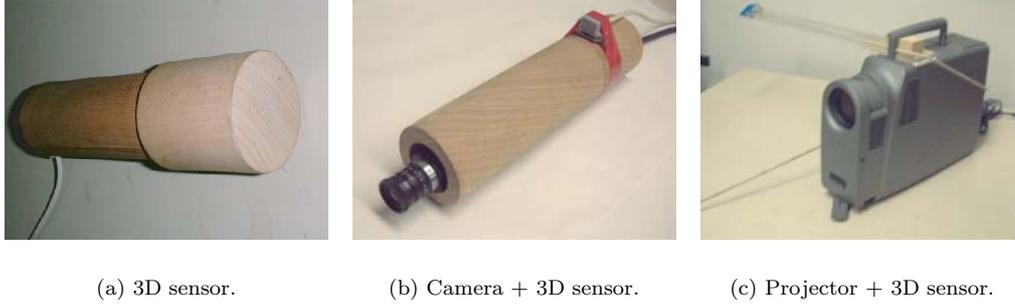If the shape of a real object can be measured in ad-

(a) 3D sensor.　　　　　(b) Camera + 3D sensor.　　　　　(c) Projector + 3D sensor.

2　3
Fig. 2　Three types of Virtual Light devices.

vance, it is easy to display its shadow interactively (see Fig. 1(a) and Sect. 5. 1). Research up to now has for the most part dealt with static or rigid objects in the real world. Here, we also investigate real-time projection onto the virtual world of a real object's shadow that is non-rigid and changing over time (see Fig. 1(b) and Sect. 5. 2) [7].

Also, in research on the superimposing of virtual objects on images of the real world, State et al. [10] have proposed a method for superimposing the shadow of a virtual object (a knot) onto a real object (a sculpture) by utilizing a polygonal model of the sculpture. This means that the shape of the real object must be measured in advance. Sato et al. [9] have proposed a technique for estimating the lighting environment of the real world and then rendering the shading and shadowing of a virtual object to conform to that environment. This technique is especially effective for synthesizing high-quality images. This is, however, a method for image synthesis, so we can see the result just in a synthetic image. We go beyond image synthesis and consider the actual projection of a virtual object's shadow onto the real world whose shape is not measured (see Fig. 1(d) and Sect. 5. 4). In this implementation, we can see just the shadow, and we need an additional display (See-through HMD) to see the virtual object itself.

3.　Virtual Light

An ordinary flashlight emits light on its own and can therefore be used to illuminate space as desired. If, however, we were to know the position and direction of a flashlight that a user is holding and then project the corresponding image on a screen, the user could experience the same visual effect as a flashlight even if the flashlight is not actually emitting light. Here, the interface that provides the same or more versatile visual effects of a flashlight, regardless of whether the flashlight is emitting light, is called "Virtual Light."

Examples of Virtual Light devices are shown in Fig. 2. Example (a) is a 3D sensor for measuring 3D position and direction that a user can hold like a flashlight. Example (b) embeds a CCD camera in the 3D sensor and is used for projecting a real object's shadow that changes over time onto the virtual world (see Fig. 1(b) and Sect. 5. 2). Example (c), on the other hand, uses a projector instead of a camera to project a virtual object's shadow while illuminating the real world with light emitted from the projector itself (see Fig. 1(d) and Sect. 5. 4). For our purpose, it is obvious that we need a smaller projector than example (c). We are planning to develop a new projection system as compact as a flashlight in the future work.

4.　Virtual Shadow

A shadow synthesized and displayed on the basis of Virtual Light is called a "Virtual Shadow." In this regard, an illuminated object and the world onto which the shadow of that object is projected may be either real or virtual. We therefore can consider four types of Virtual Shadows as summarized in Table 1.

A real object means an object which actually exists in front of a user. On the other hand, a virtual object means an object displayed on a screen. In this context, a picture of a real object displayed on a screen is regarded as a virtual object.

4. 1　RR Shadow

The RR (Real to Real) Shadow includes the case where a real object in the real world is illuminated by an ordinary flashlight. There is nothing special
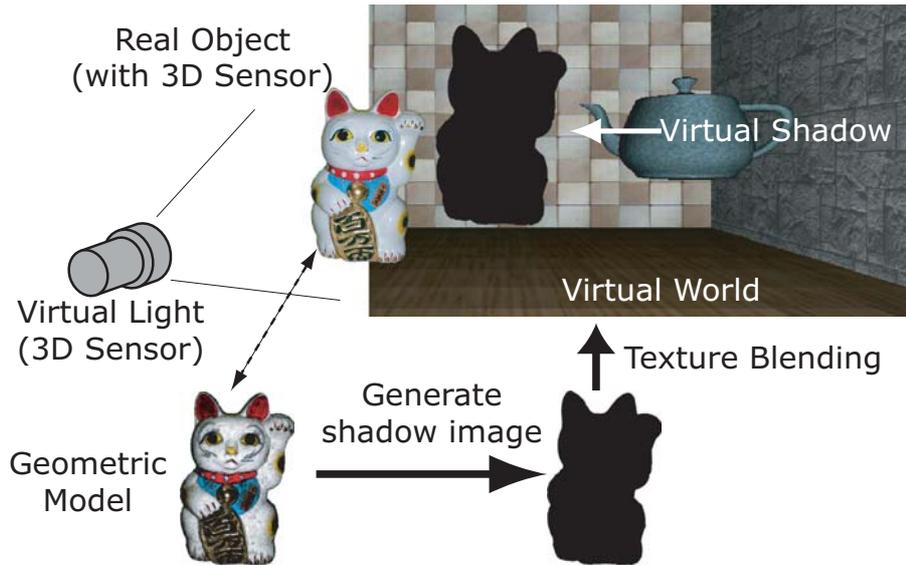
3　　　　　　　　RV Shadow
Fig. 3　Mechanism of RV Shadow for rigid objects.

1　4
Table 1　Four types of Virtual Shadows.

	Real World	Virtual World
	RR Shadow	RV Shadow
Real Object	Shadows of real objects onto the real world	Shadows of real objects onto the virtual world
	VR Shadow	VV Shadow
Virtual Object	Shadows of virtual objects onto the real world	Shadows of virtual objects onto the virtual world

in this case. On the other hand, as described by Chikamori et al. [1], a new form of visual expression can be achieved by synthesizing and displaying an unrealistic shadow of a real object onto the real world.

4. 2　RV Shadow

The RV (Real to Virtual) Shadow is a technique that modifies a real object's shadow in the virtual world according to the way in which the user is manipulating the real object. Here, we divide real objects into rigid and non-rigid ones and implement separate systems for these two cases (see Figs. 1(a) and (b), and Sects.5.1 and 5.2). In both cases, the shape of the real object is measured and utilized for projecting its shadow onto the virtual world whose shape is given.

4. 3　VV Shadow

The VV (Virtual to Virtual) Shadow corresponds to the projection of a virtual object's shadow onto the virtual world as achieved by traditional computer graphics (CG). There is nothing special in the case

where shape of both the virtual object and the virtual world are given. In this paper, in order to treat the virtual object whose shape is unkown, we apply the image-based method (see Fig. 1(c) and Sect.5.3).

4. 4　VR Shadow

The VR (Virtual to Real) Shadow is a technique for achieving a seamless Mixed Reality Environment by projecting the shadow of a virtual object onto the real world. We consider the case where the shape of the virtual object is given. In this paper, our aim is not to synthesize images, but to actually project a virtual object's shadow onto the real world whose shape is unknown (see Fig. 1(d) and Sect.5.4).

5.　Experiments

We introduce four examples of implementing Virtual Light Interaction. The corresponding experimental setups are shown in Fig. 1 and described below in order.

5. 1　RV Shadow for rigid objects

Figure 3 illustrates the mechanism behind RV Shadow for rigid objects shown in Fig. 1(a).

To begin with, we measure the 3D shape of the real object (here, a "manekineko," a popular figure of a cat in Japan displayed in shops to bring good luck) by a range finder and create a geometric model. Then, to interact, we use the Virtual Light shown in Fig. 2(a) and attach a 3D sensor to the cat figure. Finally, we synthesize an appropriate shadow image corresponding to the cat's position and orientation by using its geometric model and project that
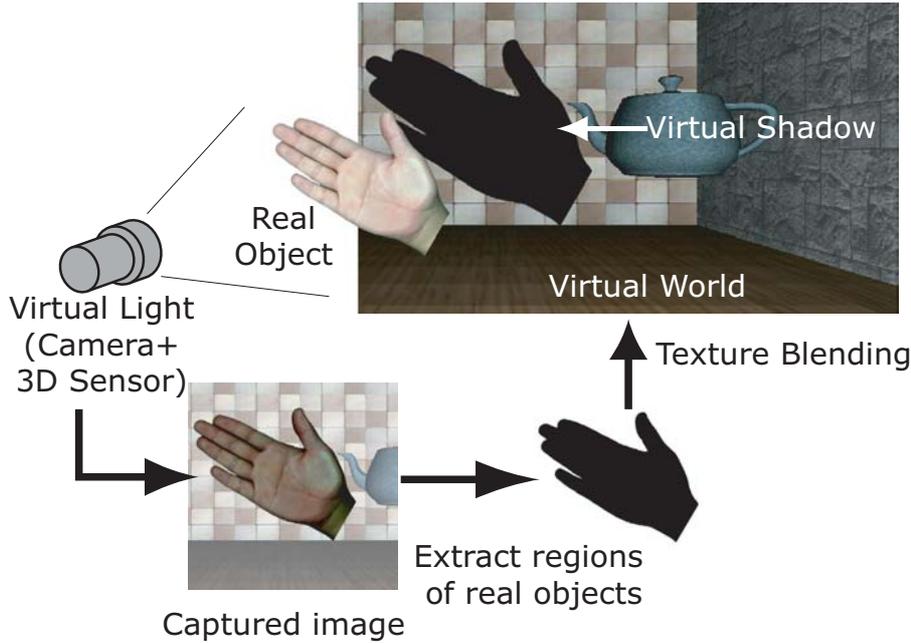
4   RV Shadow

Fig. 4   Mechanism of RV Shadow for non-rigid objects.

shadow onto the virtual world.

To be more specific, we synthesize the image that would result by looking at the geometric model positioned at the cat's position from the position of Virtual Light and use the silhouette of that image as a shadow image. This shadow image is then subjected to an α blending with texture at appropriate positions in the virtual world.

As described above, image synthesis in this example can be achieved by conventional CG technology using geometric models. The technique described here, however, contributes to the realization of a seamless Mixed Reality Environment by creating the visual effect whereby the shadow of a real object appears to be projected onto the virtual world.

5.2   RV Shadow for non-rigid objects

Figure 4 illustrates the mechanism behind RV Shadow for dynamic objects shown in Fig. 1(b). When a real object can change form in various ways like a shadow puppet, its shape obviously cannot be measured beforehand. For this reason, we use the Virtual Light shown in Fig. 2(b) and use the camera to capture such changes in the real world and reflect them in the virtual world.

In this case, the system picks up both the real world and the virtual world displayed on the screen from the position of the Virtual Light. This requires real objects to be extracted from the image picked

up by the camera. Let us consider the case where the virtual world is displayed on a large screen inside a dark room. This is very common for Immersive Projection Technology (IPT) systems such as CAVE, CABIN and COSMOS. In this case, the real object is very dark since it is illuminated just by the projected screen image. So, we can assume that bright sections within camera input correspond to the virtual world and dark sections to real objects. This assumption is useful for extracting real objects from the image picked up by the camera.

Figure 5 shows the specific procedure for image synthesis. We first synthesize a shaded view $S(x,y)$ of the virtual world illuminated from the position of the Virtual Light and then synthesize a shadow map $M(x,y)$ that describes how the shadows of real and virtual objects appear. Here, we can synthesize a virtual-world image $I(x,y)$ by multiplying the color value $S(x,y)$ and the shadow factor $M(x,y)$ together for each pixel (x,y): $I(x,y) = M(x,y)S(x,y)$, where $0 < M(x,y) \leqq 1$. This method, however, would cause very dark sections throughout the screen. This is not good for the extraction process mentioned above. We therefore synthesize an ambient view $A(x,y)$ assuming the existence of ambient light having a fixed level of brightness. Then, by subjecting these three images to α blending, we synthesize a virtual world image with brightness at that level or

(a) Shaded view $S(x, y)$.

(b) Shadow map $M(x, y)$.
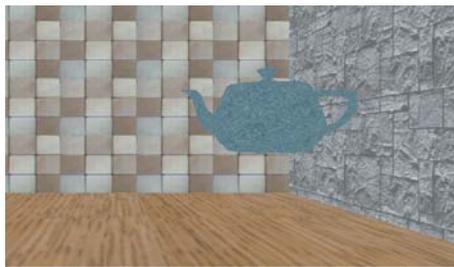
(c) Ambient view $A(x, y)$.

(d) Final result $I(x, y)$.

5　　　　RV Shadow
Fig. 5　Process of synthesizing RV Shadow for non-rigid objects.

6　Shadow map $M(x, y)$
Fig. 6　Synthesis of shadow map $M(x, y)$.

higher so that real objects picked up by the camera can now be extracted. This blending process can be expressed as follows:

$$I(x, y) = (1 - t)M(x, y)S(x, y) + tA(x, y). \quad (1)$$
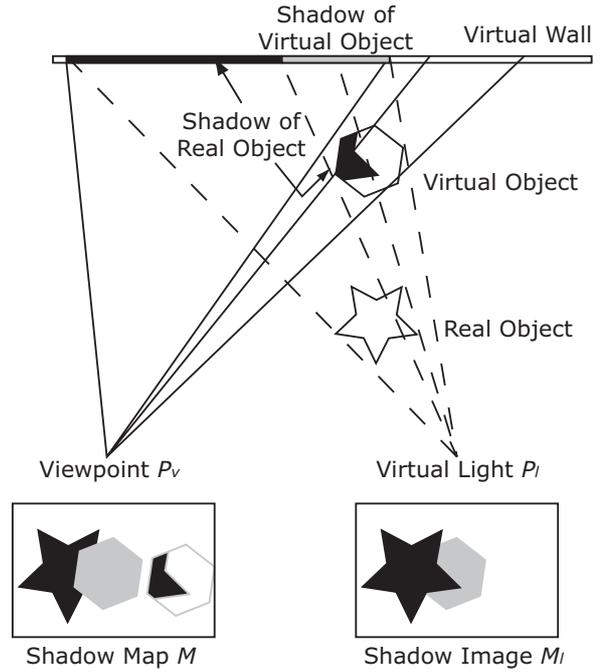
In this experiment, $t = 0.25$.

The procedure for synthesizing the shadow map $M(x, y)$ is as follows (see Fig. 6). An image $M_l(x_l, y_l)$ in the figure describes the shapes of real objects and virtual objects viewed from the position of the Virtual Light P_l. This means that $M_l(x_l, y_l)$ represents the shape of shadows that should be projected from P_l. Since our viewpoint P_v is apart from P_l, we would see a shadow map $M(x, y)$ that is a distorted and displaced version of $M_l(x_l, y_l)$. In order to synthesize $M(x, y)$ from $M_l(x_l, y_l)$, we first project and map the shadow image $M_l(x_l, y_l)$ onto the virtual world model. Next, while looking from P_v at this virtual world on which only shadows have been projected from P_l, we synthesize the shadow map $M(x, y)$.

The following describes the method for synthesizing this shadow image $M_l(x_l, y_l)$ from camera input. Denoting the brightness of each pixel of camera input as $l(x_l, y_l)$, the value of $M_l(x_l, y_l)$ can be calculated by the following expression in terms of threshold values t_1 and t_2.

$$M_l = \begin{cases} M_0 & (0 \leqq l < t_1) \\ M_0 + (1 - M_0)\frac{t_1 - l}{t_1 - t_2} & (t_1 \leqq l < t_2) \quad (2) \\ 1 & (t_2 \leqq l) \end{cases}$$

Here, M_0 is the minimum value of $M_l(x_l, y_l)$ and corresponds to a real object. When the pixel value of camera input $l(x_l, y_l)$ is small or dark enough, $M_l(x_l, y_l) = M_0$. On the other hand, $M_l(x_l, y_l) = 1$
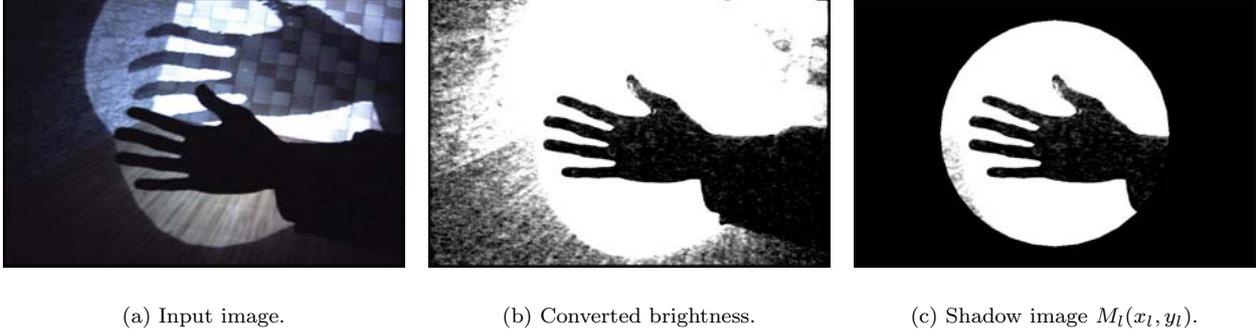
(a) Input image.   (b) Converted brightness.   (c) Shadow image $M_l(x_l, y_l)$.

7
Fig. 7   Synthesizing a shadow image of a non-rigid real object.
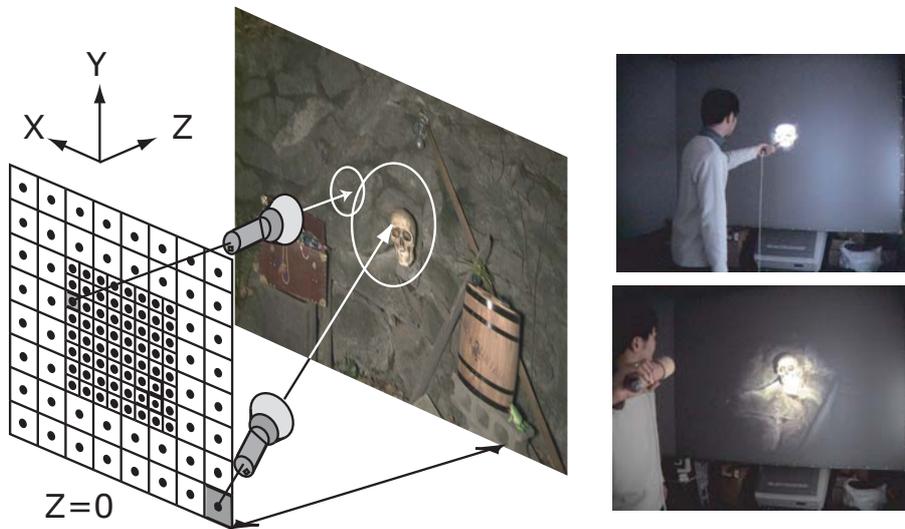

8   VV Shadow   .
Fig. 8   Mechanism of Image-based VV Shadow.

means that the pixel $l(x_l, y_l)$ corresponds to the virtual world. For ambiguous regions such as $t_1 \leqq l(x_l, y_l) < t_2$, intermediate value is assigned, that is, $M_0 < M_l(x_l, y_l) < 1$. Compared to the binary extraction technique, this technique produces a blur effect at shadow boundaries making for good visual results.

Figure 7 shows the experimental results of synthesizing a shadow image $M_l(x_l, y_l)$. The image captured by the Virtual Light is shown in (a), the result of applying Eq.(2) is in (b), and the spotlight effect is added in (c). Experimental results show that the proposed algorithm is robust when the virtual world is displayed on a large screen inside a dark room.

5.3   Image-based VV Shadow

If the shape of a virtual object is known, it is easy to project its shadow onto the virtual world. Let's consider the case where we can not utilize the shape

data of objects. In Fig. 1(c), images taken beforehand under a variety of lighting conditions are selectively displayed according to the position and direction of Virtual Light shown in Fig. 2(a).

Consider a grid on the XY-plane ($Z = 0$) as shown in Fig. 8. While moving the position of a point light source from one grid cell to another in order, we take one real-world photograph at each position. Throughout this process, the position of the camera is completely fixed. In the experiment, moreover, the light source is moved at finer intervals in the center portion of the grid than on the outside resulting in more images for the center area. This makes it possible to move the Virtual Light over a wide range and also to observe subtle changes in shading and shadowing in the center area. This capturing strategy is designed for interaction purpose, and has not been considered in previous researches.
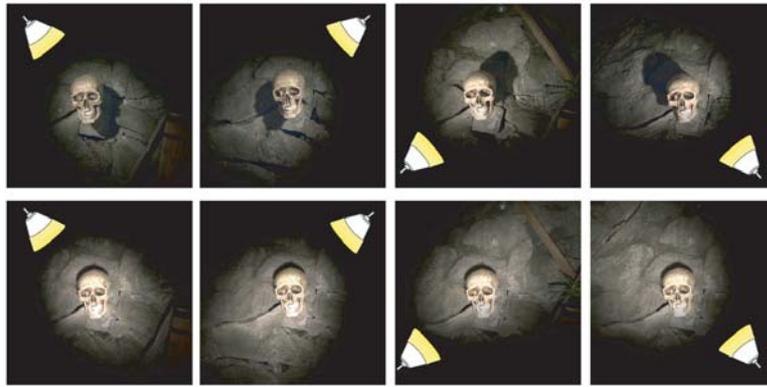
9   :   :
Fig. 9   Effects of image-based approach (upper: proposed technique, lower: conventional technique).

Specifically, the distance between the real object and camera was 1.8 m and the light-source interval in the center and outside areas was 9 cm and 18 cm, respectively. We moved the light source over an 8×8 grid covering a 126 cm square area and similarly took 8×8 photographs in the 63 cm square center area.

When the user brings the position of the Virtual Light to the $Z = 0$ plane, an appropriate image can be displayed by determining the nearest light-source position and selecting the image corresponding to that position. However, if the position of the Virtual Light is not at the $Z = 0$ plane, another means of image display must be found. We therefore decided to synthesize images under lighting conditions for which no photographs were taken, by illuminating smaller areas as the user approaches the screen and illuminating larger areas as the user moves away from the screen, as shown on the right in Fig. 8.

Furthermore, to select images in this case, we considered not only the position of the Virtual Light but its direction as well. In this process, we draw a straight line from the position of the Virtual Light along the direction to the screen, calculate coordinates (X, Y) where this line would intersect the $Z = 0$ plane, and select the picture corresponding to the light-source position nearest that point, as shown on the left in Fig. 8. In addition, by performing an α blending between multiple images, a smooth change in images could be achieved as the Virtual Light moves.

Figure 9 shows examples of images displayed in the above manner. The bottom row shows results obtained by the conventional technique that simply modifies the illuminated area with respect to one im-

age, while the top row shows results obtained by the proposed technique that uses multiple images. When interacting, the user receives a rather "flat" impression with the technique using only one image, but has a stronger sensation of a "solid" object with the proposed technique.

5. 4   VR Shadow

Figure 10 shows the mechanism behind VR Shadow shown in Fig. 1(d). In this case, we use the Virtual Light shown in Fig. 2(c) to project the shadow of a virtual object onto the real world while actually illuminating the real world. The silhouette of the virtual object (geometric model) seen from the position of the Virtual Light becomes the shadow image that must be projected. The user must wear a See-Through HMD or the like to view the virtual object, since just the shadow appears in the real world.

Figure 11(a) shows an example of normal lighting in the real world and Fig. 11(b) shows the same picture with a virtual object (a teapot) superimposed. Figure 11(c), moreover, shows the projection of a virtual object's shadow onto the real world. Note that a user not wearing a See-Through HMD would simply see a shadow appears from nothing. This user, however, would be able to estimate the position and shape of the virtual object by moving the Virtual Light and changing the shadow interactively. If the virtual object is now superimposed, we get the scene shown in Fig. 11(d). Compared to Fig. 11(b), the virtual object has merged well with the real world.

Figure 12 shows two other examples, which demonstrate that a visually appealing shadow can be projected even in a complex real world. This is because the proposed technique can project a shadow image
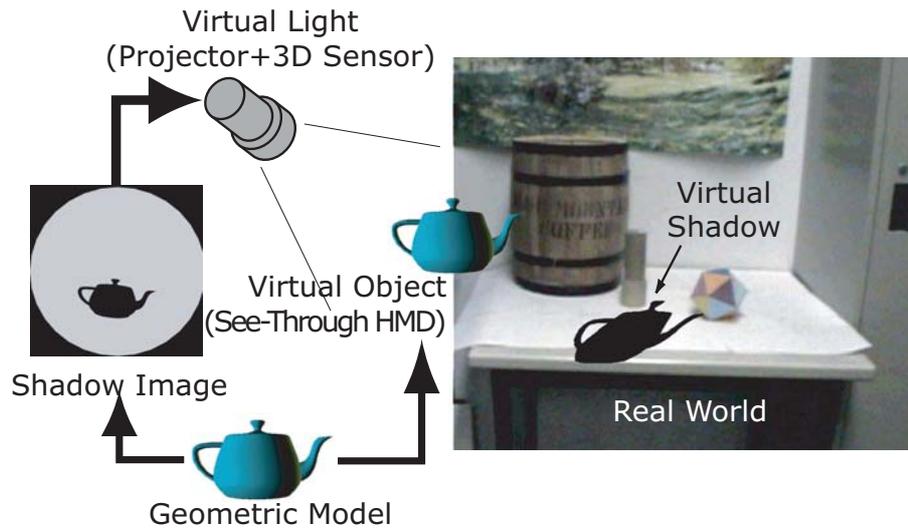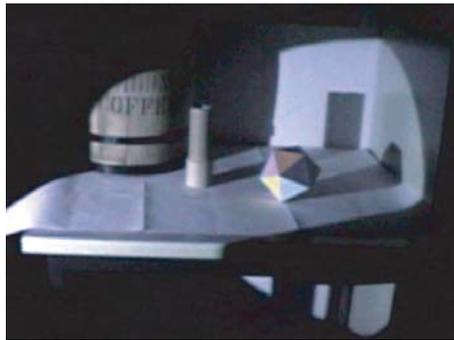
10 VR Shadow .
Fig. 10 Mechanism of VR Shadow.

(a) Normal lighting. Everyting in this scene is real.

(b) A teapot is superimposed as a virtual object. But, it does not have any shadow.

(c) A VR Shadow is projected onto the real world. This is a real picture.

(d) A teapot is superimposed. We can see that the projected VR shadow matches the superimposed virtual object.

11 VR Shadow
Fig. 11 Process of VR Shadow.

12   VR Shadow
Fig. 12   Examples of VR Shadow.

in a relatively simple manner based only on the positional relationship between the virtual object and Virtual Light; it does not depend on the shape of the real world.

6.   Conclusions

We have proposed the concepts of Virtual Light and Virtual Shadow with the aim of achieving a Mixed Reality Environment focused on shadows, and implemented four specific applications. Each of these application examples raises issues for study, as described below.

In RV Shadow for rigid objects, the system projects onto the virtual world only the shadow of a cat figure based on its geometric model. The shadow of the hand handling the cat is not drawn and the shadow makes it look as if the cat is floating in midair. This can be a useful technique for distinguishing between effective and non-effective objects in both the real and virtual worlds. New means of expression are expected to appear in a Mixed Reality World that merges parts of the real and parts of the virtual.

In RV Shadow for non-rigid objects, we assumed that interaction is performed in front of a bright screen in a dark room. It should also be possible, though, to implement this technique in various environments.

In Image-based VV Shadow, the system takes images by a fixed camera. This means that the user can not move his/her viewpoint freely. There is also room for more study on the arrangement of light sources and the taking of pictures.

In VR Shadow, a flashlight-like compact projector is required. The proposed method can be extended to embed various kinds of information in a shadow projected onto the real world.

The experimental results presented are just examples – various types of applications can be envisioned. Combining these examples could be a challenging research topic. The key technology for this purpose might be the Virtual Light interface that can capture and project images simultaneously. In addition to the three types of Virtual Light introduced in Fig. 2, range finders and infrared cameras, infrared floodlights, and omni-directional cameras can also be considered for this role. The development of new types of Virtual Light should lead to new and various kinds of interaction. And, in addition to accurate shadow rendering, we can consider the implementation of various kinds of unrealistic effects such as Magic Light [2] and Kage [1] and expansion to new methods of interaction and communication.

[1]   M. Chikamori and K. Kunoh. Kage. In *SIGGRAPH 98: Electronic Art and Animation Catalog*, page 14. ACM, 1998.

[2]   S. Inoue. Internal representation by the magic light. In *SIGGRAPH 98: Conference Abstracts and Applications*, page 288. ACM, 1998.

[3]   A. Katayama, Y. Sakagawa, and H. Tamura. A method of shading and shadowing in image-based rendering. In *Int. Conf. on Image Process. 98*, volume 3, pages 26–30. IEEE, 1998.

[4]   B. Lutz and M. Weintke. Virtual dunhuang art cave: A cave within a CAVE. In *Eurographics 99*, 1999.

[5]   A. Mimura, T. Nitta, T. Naemura, and H. Ha-

rashima. Virtual lighting interaction in image-based rendering. In *IEICE Annual Conf.*, pages A–16–18, Mar. 2000. (in Japanese).

[6] K. Nishino, Y. Sato, and K. Ikeuchi. Appearance compression and synthesis based on 3d model for mixed reality. In *Int. Conf. on Computer Vision 99*, pages 38–45. IEEE, 1999.

[7] T. Nitta, T. Naemura, and H. Harashima. Interaction of shadows in mixed reality environment. In *VRSJ Annual Conf.*, pages 459 – 462, Sept. 2000. (in Japanese).

[8] Y. Ohta and H. Tamura. *Mixed Reality – Merging Real and Virtual Worlds*. Springer Verlag, Berlin, 1999.

[9] I. Sato, Y. Sato, and K. Ikeuchi. Acquiring a radiance distribution to superimpose virtual objects onto a real scene. *IEEE Trans. Visualization and Computer Graphics*, 5(1):1–12, Jan. 1999.

[10] A. State, G. Hirota, D. T. Chen, W. F. Garrett, and M. A. Livingston. Superior augmented reality registration by integrating landmark tracking and magnetic tracking. In *SIGGRAPH 96*, pages 429 – 438. ACM, 1996.

[11] T. Wong, P. Heng, and C. Fu. Interactive relighting of panoramas. *IEEE Computer Graphics and Applications*, 21(2):32–41, Mar. 2001.

[12] T. Wong, P. Heng, S. Or, and W. Ng. Image-based rendering with controllable illumination. In *Eighth Eurographics Workshop on Rendering*, pages 13–22, 1997.

(2002　1　31　　)

Takeshi Naemura

received his BE, ME, and PhD in electronic engineering from the University of Tokyo, Tokyo, Japan, in 1992, 1994, and 1997, respectively. He is an associate professor of the Interfaculty Initiative in Information Studies at the University of Tokyo. He is currently a visiting assistant professor in computer graphics at Stanford University and supported by "Japan Society for Promotion of Science (JSPS) Postdoctoral Fellowships for Research Abroad." His research interests include mixed reality, augmented spatial communications, image-based rendering and human interfaces.

Takuya Nitta

received his BE and ME in information and communication engineering from the University of Tokyo, Tokyo, Japan, in 1999 and 2001, respectively. He is working for FUJITSU Ltd.

Atsushi Mimura

received his BE and ME in information and communication engineering from the University of Tokyo, Tokyo, Japan in 2000 and 2002, respectively. His research interests include mixed reality, evolutionary computation and bioinformatics. He is working for SONY Corp.

Hiroshi Harashima

received the B.E., M.E. and Dr.E. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1968, 1970 and 1973, respectively. He is a professor of the Interfaculty Initiative in Information Studies at the University of Tokyo. His research interests include communication theory, image coding and processing, and human communication engineering. He is the president of the Virtual Reality Society of Japan.